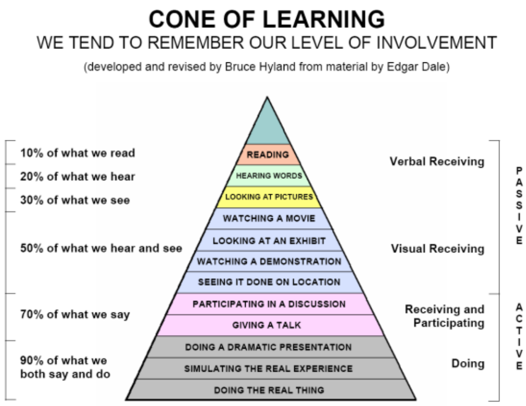


Feedback requested

- Overhead transparencies – keep or convert to PowerPoint?
- Section handout
  - Short outline
  - Short outline with all visuals
  - Expanded lecture notes

Sampling Design Basics

Objectives:

- Understand how attention to basic principals of sampling design can improve the outcome of monitoring projects.
- Identify: population, sampling unit, sample.
- List 3 types of non-sampling errors.
- Be able to calculate a 95% confidence interval from actual sampling data.
- List 3 ways to increase the Power of a monitoring study.

Topic Outline

A. Example of a failed monitoring project
B. Introduction to sampling
  1. Definition of sampling
  2. Why sample?
C. Key terms, important principals:
  1. Populations and samples
  2. Population parameters vs. sample statistics
  3. Accuracy vs. Precision
  4. Standard Error
  5. Confidence Intervals
  6. Finite vs. Infinite Populations
  7. Sampling vs. nonsampling errors
  8. False-change Errors, Missed-change Errors, Power, and Minimum Detectable Changes

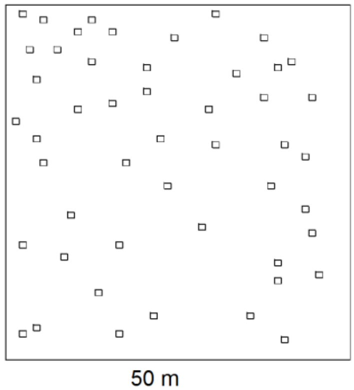
Topic Outline continued

Exercises

S1: Sampling a clumped population
S2: Identifying populations, sampling units, samples
S3: Calculating confidence intervals
S3.5: Power comparisons

Monitoring changes in *Lomatium Cookii*

Lomatium cookii macroplot at the Agate Desert



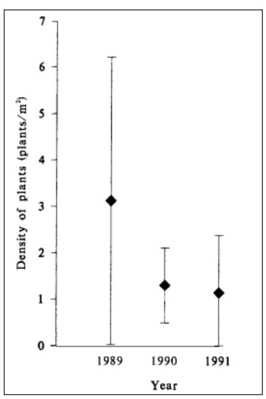


Figure 1. Density of *Lomatium cookii* in Macroplot 2 at the Agate Desert, 1989-91. Bars represent 95% confidence intervals. *Lomatium cookii* were counted in 50 $1m^2$ plots each year.

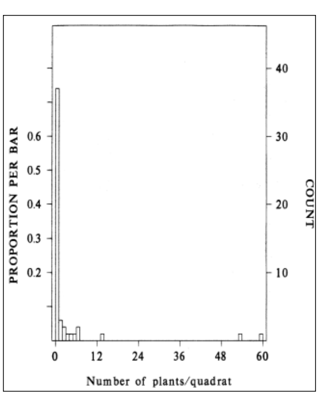
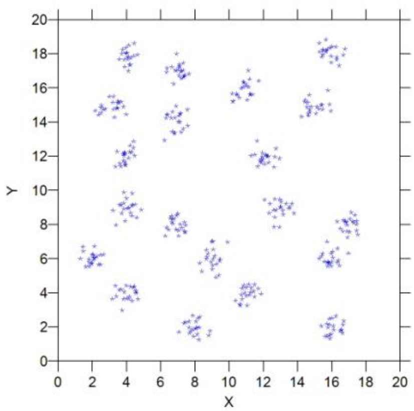


Figure 2. Frequency histogram of number of *Lomatium cookii* plants per $1m^2$ quadrant in macroplot 2 at the Agate Desert in 1989. (n=50 quadrants; sum=156 plants; mean # plants/ quadrat=3.12; sd=11.17; 37 quadrats with no plants; 13 with plants; 3-1 plant, 2-2 plants, and 1 quadrat with each of the following counts: 3, 4, 5, 13, 53, 59).

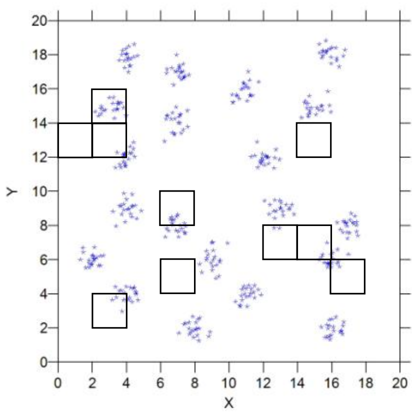
Definition of sampling:

The act or process of selecting a <u>part</u> of something with the intent of showing the quality, style, or nature of the <u>whole</u>.

400 plant population



A random arrangement of 10 - 2m x 2m quadrats positioned within a 400-plant population.



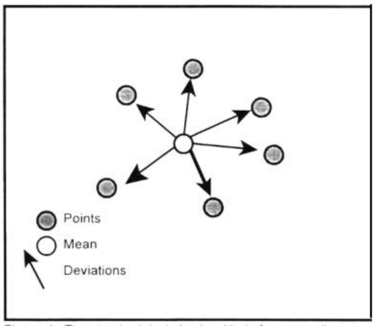
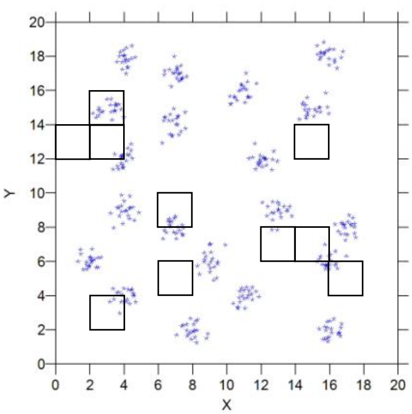


Figure 4. The standard deviation is a kind of average distance between the observations and the mean of all the observations.

A random arrangement of 10 - 2m x 2m quadrats positioned within a 400-plant population.



Sample information		
Coordinates		# of plants
X	Y	
2	2	4
6	4	0
16	4	3
12	6	2
14	6	5
6	8	10
0	12	0
2	12	6
14	12	0
2	14	20

Sample statistics (n=10)

Mean # plants/quadrat

$\bar{x} = 5.0$

Standard deviation:

$s = 6.146$

Population estimate

Est. pop. size = 500 plants

95% CI = ± 361 plants

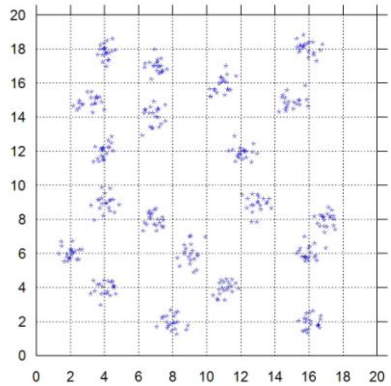
Sample statistics for the 400-plant population.

Population parameters

Tot. pop. size:
400 plants

Mean # plants / quadrat (2m x 2m):
$\mu = 4$

Standard deviation:
$\sigma = 5.005$

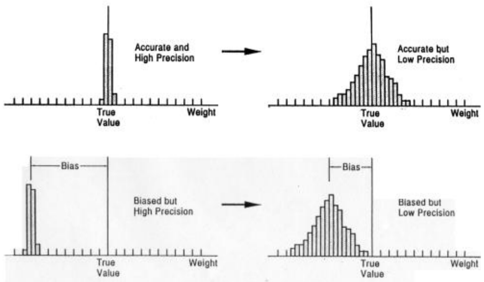


Sample information			Population parameters
Coordinates		# of plants	Tot. pop. size: 400 plants
			Mean # plants/quadrat:
X	Y		$\mu = 4$
2	2	4	Standard deviation:
6	4	0	$\sigma = 5.005$
16	4	3	**Sample statistics** (n=10)
12	6	2	Mean # plants/quadrat
14	6	5	$\bar{x} = 5.0$
6	8	10	Standard deviation:
0	12	0	$s = 6.146$
2	12	6	**Population estimate**
14	12	0	Est. pop. size = 500 plants
2	14	20	
			95% CI = ± 361 plants

Population parameters and sample statistics for the 400-plant population.

Accuracy is the closeness of a measured or computed value to its true value

Precision is the closeness of repeated measurements of the same quantity



An illustration of accuracy and precision in ecological measurements. In each case, a series of repeated measurements are taken on a single item, e.g. weight of a single fish specimen. From Krebs, C.J. 1989. Ecological Monitoring. Harper Collins, New York.

Sample with High Precision	Sample with Low Precision
9	2
10	10
14	21
Mean = 11	Mean = 11
Standard Deviation = 2.65	Standard Deviation = 9.54

Formula for standard error

$$SE = \frac{s}{\sqrt{n}}$$

Where:

s = sample standard deviation

n = sample size

Standard formula for a confidence interval

$$\text{C.I.}_{\text{half-width}} = \text{SE} \times t_{\text{value}}$$

Critical t-values for several levels of confidence (for 2-sided confidence intervals).				
degrees of freedom	**80%**	**90%**	**95%**	**99%**
1	3.078	6.314	12.706	63.656
2	1.886	2.920	4.303	9.925
3	1.638	2.353	3.182	5.841
4	1.533	2.132	2.776	4.604
5	1.476	2.015	2.571	4.032
6	1.440	1.943	2.447	3.707
7	1.415	1.895	2.365	3.499
8	1.397	1.860	2.306	3.355
9	1.383	1.833	2.262	3.250
10	1.372	1.812	2.228	3.169
11	1.363	1.796	2.201	3.106
12	1.356	1.782	2.179	3.055
13	1.350	1.771	2.160	3.012
14	1.345	1.761	2.145	2.977
15	1.341	1.753	2.131	2.947

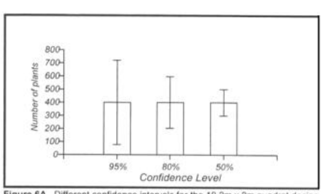


Figure 6A. Different confidence intervals for the 10 2m x 2m quadrat design

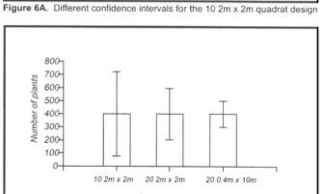
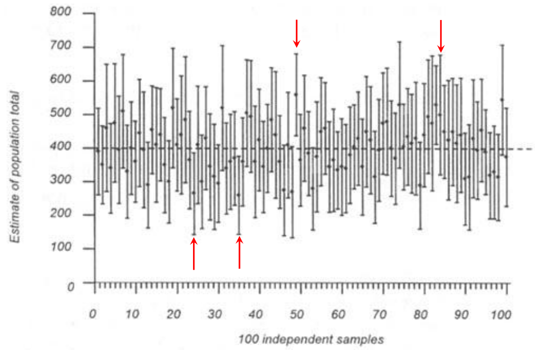


Figure 6B. 95% confidence levels for 3 different sampling designs.

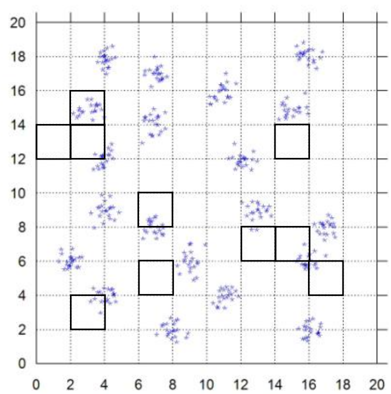


Use the Finite Population Correction when sampling with quadrats and > 5% of all possible quadrats

n=10

N=100

10% of all quadrats sampled

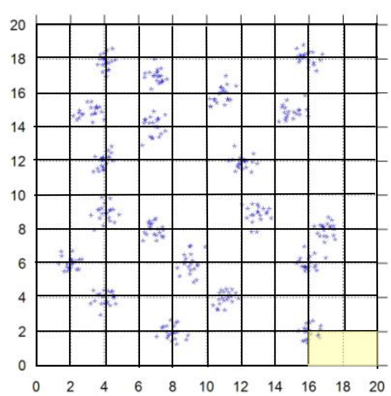


Finite Population Correction "rewards" you for knowing more about the population

n=98

N=100

98% of all quadrats sampled



Formula for the finite population correction (FPC)

$$FPC = \sqrt{\frac{N - n}{N}}$$

Where:
N = The total number of potential quadrat positions
n = The number of quadrats sampled

Example of calculating an FPC

- Total population area = 20m x 50m macroplot (1000 m^2)
- Size of individual quadrat = 10 m^2
- Sample size (n) = 30 quadrats

$$N = \frac{1000\ \text{m}^2}{10\ \text{m}^2} = 100$$

$$FPC = \sqrt{\frac{N - n}{N}} \qquad 0.83 = \sqrt{\frac{100 - 30}{100}}$$

Standard formula for a confidence interval when sampling from a finite population

$$\text{C.I.}_{\text{half-width}} = \text{SE} \times t_{\text{value}} \times \text{FPC}$$

Non-Sampling Errors

- Biased selection rules
- Unrealistic or inappropriate techniques
- Sloppy field work
- Transcription and recording errors
- Inconsistent species identification

Sampling Errors

- The difference between a sample-based estimate and the true population
- Errors resulting from chance – an inevitable consequence of the sampling process

Sampling Error

Sample statistics (n=10)

Mean = 0.8 plants / quadrat:

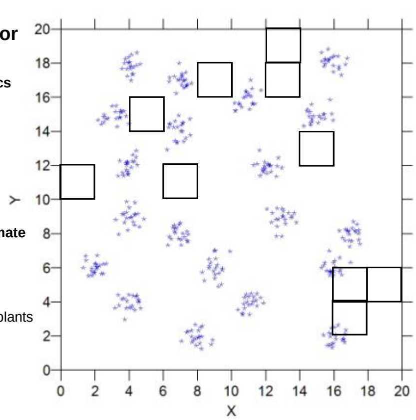
s = 1.75

Population estimate

Est. pop. size =

80 plants

95% CI = ± 119 plants



Sampling Error

Sample statistics
(n=10)

Mean = 9.6 plants / quadrat:

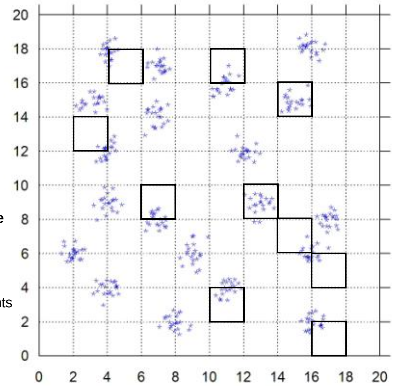
s = 5.58

Population estimate

Est. pop. size =

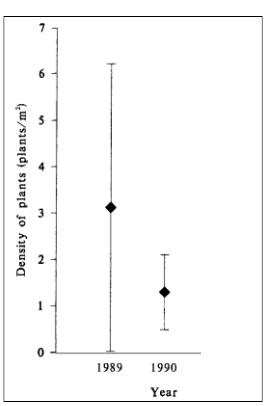
960 plants

95% CI = ± 379 plants



Topic Outline

A. Example of a failed monitoring project
B. Introduction to sampling
  1. Definition of sampling
  2. Why sample?
C. Key terms, important principals:
  1. Populations and samples
  2. Population parameters vs. sample statistics
  3. Accuracy vs. Precision
  4. Standard Error
  5. Confidence Intervals
  6. Finite vs. Infinite Populations
  7. Sampling vs. nonsampling errors
  8. False-change Errors, Missed-change Errors, Power, and Minimum Detectable Changes



Lomatium cookii
1989-1990

- 63% decline in sample mean
- But did a change really take place?

Monitoring for change: possible errors

	No real change has taken place	There has been a real change
Monitoring system detects a change	**False-change Error (Type I) α**	**No Error (Power) 1-β**
Monitoring system detects no change	**No Error (1-α)**	**Missed-change Error (Type II) β**

Monitoring for change: possible errors

	No real change has taken place	There has been a real change
Monitoring system detects a change	**False-change Error (Type I) α**	**No Error (Power) 1-β**
Monitoring system detects no change	**No Error (1-α)**	**Missed-change Error (Type II) β**

The origin of the 0.05 "threshold"

- R.A. Fisher (1936)
  - If P is between 0.1 and 0.9 there is certainly no reason to suspect the hypothesis tested. If it is below 0.02 it is strongly indicated that the hypothesis fails to account for the whole of the facts. We shall not often be astray if we draw a conventional line at 0.05…

Monitoring for change: possible errors

	No real change has taken place	There has been a real change
Monitoring system detects a change	**False-change Error (Type I) α**	**No Error (Power) 1-β**
Monitoring system detects no change	**No Error (1-α)**	**Missed-change Error (Type II) β**

(2) Increase Sample size

(3) Increase Minimum detectable change

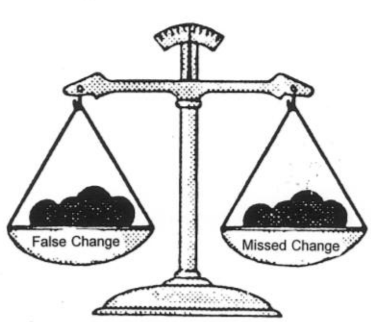
(4) Increase False-change (Type I) error

Power = a function of (s, n, MDC, and α)

(1) Decrease Standard deviation

4 ways to improve Power

Balancing The Two Kinds of Errors



Medical Field: screening patients for some lethal disease.

Null hypothesis: person does not have the disease.

Less concerned about making a false diagnosis error (Type I error, analogous to a false-change error).

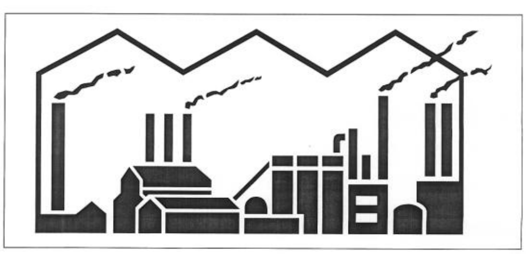
More concerned about failing to diagnose the disease (Type II error, analogous to a missed-change error).

Court of law. Null hypothesis: person is innocent

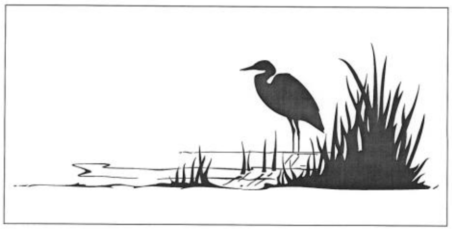
Criminal cases:
"Proof beyond reasonable doubt"
Greater chance of a guilty person going free, committing a Type II error (analogous to a missed-change error)

Civil cases:
Proof based upon the "balance of probabilities". Two types of errors closer to equality

Potential industrial pollution source. Null hypothesis: no pollution impact. Industry targets very low false-change (Type I) error rate. Industry less concerned with Power and occasional missed-change errors.



Environmental groups are more concerned about making missed-change (Type II) errors than they are false-change (Type I) errors



Uses of Power Analyses

- Prior Power Analysis (during study design)
- Post-hoc Power Analysis (for interpreting non-significant results)

Sample size

Minimum detectable change

False-change (Type I) error

Power = a function of (s, n, MDC, and α)

Standard deviation

Sample size

Minimum detectable change

False-change (Type I) error

n = a function of (s, Power, MDC, and α)

Standard deviation

Minimum detectable change

Sample size

False-change (Type I) error

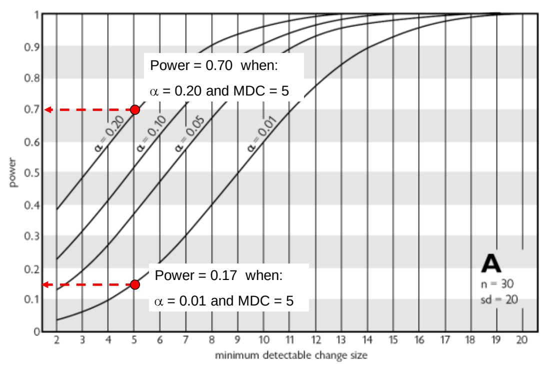
MDC = a function of (s, n, Power, and α)

Standard deviation

Prior Power Analysis on 1989 *Lomatium cookii* data

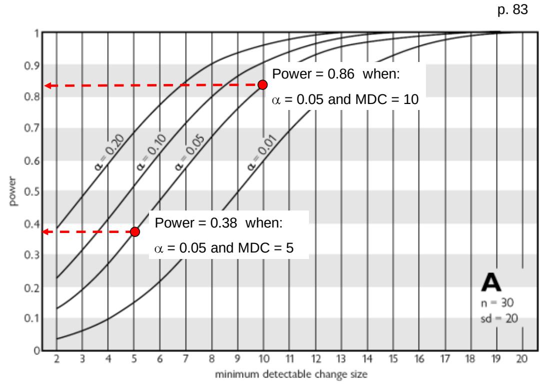
- Minimum detectable effect size with α and β = 0.10 = 200% change
- Power to detect a 50% change is only 0.18

p. 83



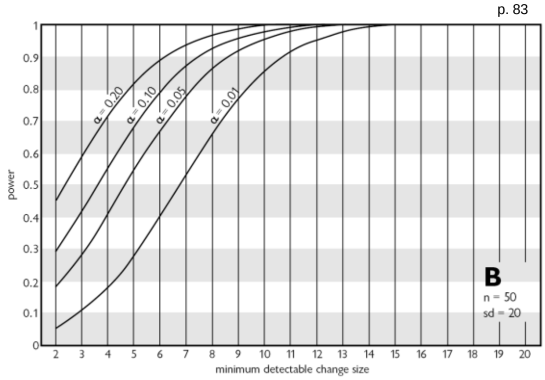
Power = 0.70 when:
α = 0.20 and MDC = 5
Power = 0.17 when:
α = 0.01 and MDC = 5
α = 0.20
α = 0.10
α = 0.05
α = 0.01
A
n = 30
sd = 20
power
minimum detectable change size

p. 83



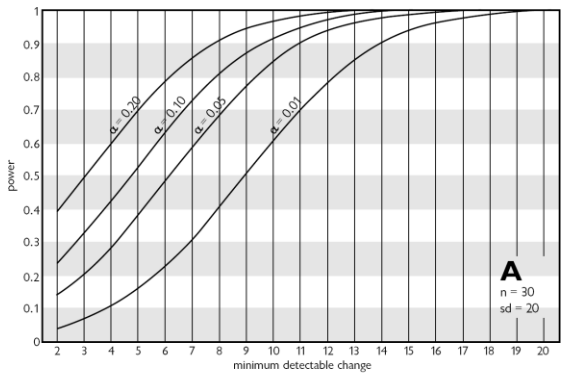
Power = 0.86 when:
α = 0.05 and MDC = 10
Power = 0.38 when:
α = 0.05 and MDC = 5
α = 0.20
α = 0.10
α = 0.05
α = 0.01
A
n = 30
sd = 20
power
minimum detectable change size

p. 83

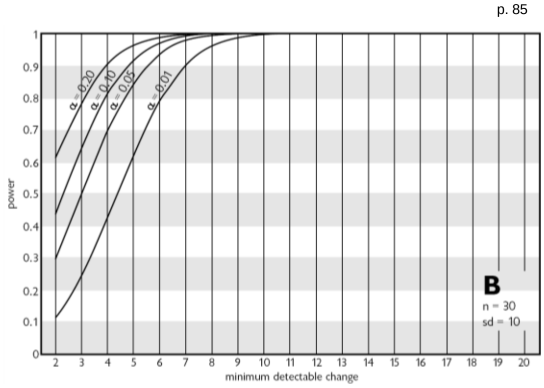


α = 0.20
α = 0.10
α = 0.05
α = 0.01
B
n = 50
sd = 20
power
minimum detectable change size

p. 85



p. 85



Exercise 3.5 – Power, False-change Error rate, MDC

Power in Figure 5.14a (p. 83)	n = 30, s = 20	
	MDC = 5	MDC = 10
α = 0.01		
α = 0.05		
α = 0.10		
α = 0.20		

Power in Figure 5.14a (p. 83)	n = 30, s = 20	
	MDC = 5	MDC = 10
α = 0.01	0.17	0.60
α = 0.05	0.38	0.85
α = 0.10	0.53	0.92
α = 0.20	0.70	0.95
Power in Figure 5.14b (p. 83)	**n = 50, s = 20**	
	MDC = 5	MDC = 10
α = 0.01	0.28	0.86
α = 0.05	0.55	0.95
α = 0.10	0.67	0.97
α = 0.20	0.83	0.99
Power in Figure 5.15b (p. 85)	**n = 30, s = 10**	
	MDC = 5	MDC = 10
α = 0.01	0.60	0.99
α = 0.05	0.85	1.0
α = 0.10	0.93	1.0
α = 0.20	0.97	1.0

Sampling Design Basics

Objectives:

- Clearly state how attention to basic principals of sampling design can improve the outcome of monitoring projects.
- Clearly state the difference between a standard deviation and a standard error
- List 3 types of non-sampling errors.
- From a brief description of a monitoring study, identify the following components: population, sampling unit, sample.
- Be able to calculate a 95% confidence interval from actual sampling data.
- List 3 ways to increase the Power of a monitoring study.